R E P O R T  R E S U M E S

ED 019 663 AL 001 141

ON THE NATURE OF SPEECH SCIENCE.

BY- PETERSON, GORDON E.

PUB DATE 67

EDRS PRICE MF-$0.25 HC-$0.36 7P.

DESCRIPTORS- *SPEECH, *ARTICULATION (SPEECH), *ACOUSTIC PHONETICS, *PHONETICS, PHONOLOGY, SPEECH IMPROVEMENT, LANGUAGE RESEARCH, RECEPTIVE LANGUAGE,

IN THIS ARTICLE THE NATURE OF THE DISCIPLINE OF SPEECH SCIENCE IS CONSIDERED AND THE VARIOUS BASIC AND APPLIED AREAS OF THE DISCIPLINE ARE DISCUSSED. THE BASIC AREAS ENCOMPASS THE VARIOUS PROCESSES OF THE PHYSIOLOGY OF SPEECH PRODUCTION, THE ACOUSTICAL CHARACTERISTICS OF SPEECH, INCLUDING THE SPEECH WAVE TYPES AND THE INFORMATION-BEARING ACOUSTIC PARAMETERS, AND THE PROCESSES OF SPEECH PERCEPTION. PHONETICS IS CONCERNED WITH THE MORE BASIC CONCEPTS IN SPEECH PHYSIOLOGY AND SPEECH ACOUSTICS. THE NOTATIONAL ASPECT OF DESCRIPTIVE PHONETICS IS BASED PRIMARILY ON SPEECH PHYSIOLOGY, AND THE CHIEF PROBLEM IN EXPERIMENTAL PHONETICS IS THE INTERPRETATION OF THE DYNAMIC PROPERTIES OF SPEECH. THE MAJOR PROCESSES OF SPEECH PRODUCTION ARE IDENTIFIED AS INITIATION, COORDINATION, INNERVATION, ARTICULATORY, AND SHAPING. THERE ARE AT LEAST THREE MAJOR AREAS OF APPLIED SPEECH SCIENCE--THE PHONOLOGICAL DESCRIPTION OF LANGUAGES, SPEECH PATHOLOGY, AND SPEECH COMMUNICATION. THE PHONOLOGICAL DESCRIPTION OF LANGUAGES IS PRIMARILY CONCERNED WITH NORMAL SPEECH AND IS BASED ON PHONETIC CONCEPTS. SUCH DESCRIPTIONS EXTEND BEYOND PHONETICS, HOWEVER, AND REQUIRE HIGHER LEVEL CONCEPTS IN PHONOLOGY, SUCH AS THE PHONEME AND THE PROSODEME. THE MANY BRANCHES OF SPEECH PATHOLOGY MAKE IT A LARGE AREA OF SCIENTIFIC RESEARCH. IT IS NOTED THAT DESCRIPTIVE PHONETICS IS INTENDED PRIMARILY FOR NORMAL SPEECH AND SHOULD BE EXTENDED TO ENCOMPASS THE DEVIATIONS COMMONLY ENCOUNTERED IN ABNORMAL SPEECH. THERE ARE MANY AREAS OF SPEECH COMMUNICATION, INCLUDING THE VOCODER TRANSMISSION OF NATURAL SPEECH AND THE DEVELOPMENT OF MACHINE AIDS TO SPOKEN LANGUAGE INSTRUCTION. A MAJOR DIVISION OF SPEECH COMMUNICATION IS SPEECH AUTOMATION, WHICH INCLUDES AUTOMATIC SPEECH SYNTHESIS AND AUTOMATIC SPEECH RECOGNITION. AMONG OTHER POTENTIAL APPLICATIONS, SPEECH AUTOMATION WOULD PROVIDE DIRECT SPEECH COMMUNICATION WITH COMPUTERS. THIS ARTICLE APPEARED IN THE ANNUAL BULLETIN FOR 1967 OF THE RESEARCH INSTITUTE OF LOGOPEDICS AND PHONIATRICS, FACULTY OF MEDICINE, UNIVERSITY OF TOKYO, HONGO, TOKYO, JAPAN. (AUTHOR/JD)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

ED019663

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

# On the Nature of Speech Science*

Gordon E. Peterson†

"PERMISSION TO REPRODUCE THIS [redacted] MATERIAL HAS BEEN GRANTED BY ICHIRO KIRIKAE

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE [redacted] OWNER."

*Abstract*

In this article the nature of the discipline of speech science is considered and the various basic and applied areas of the discipline are discussed. The basic areas encompass the various processes of the physiology of speech production, the acoustical characteristics of speech, including the speech wave types and the information-bearing acoustic parameters, and the processes of speech perception. Phonetics is concerned with the more basic concepts in speech physiology and speech acoustics. The notational aspect of descriptive phonetics is based primarily on speech physiology, and the chief problem in experimental phonetics is the interpretation of the dynamic properties of speech. The major processes of speech production are identified as initiation, coordination, innervation, articulatory, and shaping. There are at least three major areas of applied speech science: the phonological description of languages, speech pathology, and speech communication. The phonological description of languages is primarily concerned with normal speech and is based upon phonetic concepts. Such descriptions extend beyond phonetics, however, and require higher level concepts in phonology, such as the phoneme and the prosodeme. The many branches of speech pathology make it a large area of scientific research. It is noted that descriptive phonetics is intended primarily for normal speech and should be extended to encompass the deviations commonly encountered in abnormal speech. There are many areas of speech communication, including the vocoder transmission of natural speech and the development of machine aids to spoken language instruction. A major division of speech communication is speech automation, which includes automatic speech synthesis and automatic speech recognition. Among other potential applications, speech automation would provide direct speech communication with computers.

† Speech Communications Research Laboratory, Santa Barbara, California, U.S.A.

* The preparation of this paper was supported by the Directorate of Information Sciences of the United States Air Force Office of Scientific Research under Grant AF-AFOSR-1215-67.



AL 001 141

Those of us in the Speech Communications Research Laboratory warmly congratulate our friends and colleagues in the Research Institute of Logopedics and Phoniatrics on the completion of their first year of research at the Institute. The areas of investigation planned at the Institute encompass an impressive complement of important and closely related research subjects. The scientific investigations currently in progress and proposed for the Speech Science section constitute an imaginative and important set of problems at the frontiers of speech research. The problems of speech science are basic and legion, and they deserve the serious efforts of many laboratories. It is exciting to know that the scientific study of speech, the phonological description of languages, and speech pathologies will be investigated together in a single Institute in Japan. The natural emphasis on Japanese speech which one would expect to result from the location of the Institute should bring much new information about the spoken form of the Japanese language. It should be highly interesting to compare the results of this work with similar efforts on other languages. The establishment of the Research Institute of Logopedics and Phoniatrics at the University of Tokyo is an event which we welcome, and we look forward to important advances in knowledge about speech from the Institute in the years ahead.

## Introduction

In this paper I should like to consider the major branches of the field of speech science and to suggest those applied areas which I believe are strongly dependent upon basic knowledge in this field. Those who engage seriously in academic study and basic research are seeking rewards in knowledge and understanding. One soon learns that concepts in any scholarly field are ever subject to modification and that knowledge is continually expanding in both depth and breadth. Within this framework it is only natural that one should repeatedly seek to understand the nature of the discipline in which he works, to identify its major areas, and to relate its boundaries to other disciplines.

In the past, work in speech science has been associated with a variety of colleges (faculties) and departments. While there are various basic branches of the discipline, there are also applied fields which are strongly and primarily dependent upon basic knowledge in speech science.

## Basic Speech Science

There are three major basic areas of speech science, namely: speech physiology, speech acoustics, and speech perception. Because of the ready availability of the speech signal, speech acoustics is probably the most systematically developed of these three areas. In some respects this development has not been as rewarding as one might hope, however, because of the need for a substantial understanding of the physiology of speech production in order to interpret the acoustical characteristics of the speech signal. Speech perception is a potential area of study within

speech science, but this aspect of the field requires a solid foundation in speech physiology and speech acoustics. Speech physiology has been, and probably will continue to be, the most fruitful area of research in speech science. Descriptive phonetics is based on the parameters of speech physiology and is one of the most substantial and valuable tools in work on spoken languages.

The more basic aspects of the field of speech science are often denoted as the discipline of phonetics. Phonetics involves both speech physiology and speech acoustics, and is sometimes separated into physiological phonetics and acoustic phonetics. In general, phonetics is concerned with the more basic and elementary units and parameters of speech. This aspect of the field has been well established as a discipline in many universities, particularly in Europe.

As the field of linguistics developed, however, it became increasingly clear that phonetics is only the foundation of the study of spoken language, and that many concepts beyond those normally considered in the field of phonetics are required in order to describe the speech of various languages. There is the additional complication that descriptive phonetics has generally dominated the field and experimental phonetics has commanded the interest of only a limited group of scholars. The scientific development of the subdiscipline of phonetics requires a thorough and general integration of descriptive and experimental phonetics. In some instances this integration is gradually being achieved.

## Speech Physiology

Speech physiology is a particularly challenging subject because of the tremendous complexity of the mechanism of speech production. Many processes are involved in forming the speech signal. These include: (1) the *initiation process* with which linguistic messages are formulated at a relatively high level within the central nervous system; (2) the *coordination process* whereby neural instructions are organized and integrated; (3) the *innervation process* which provides direct motor control of the speech musculature; (4) the *articulatory process* in which the relatively independent actions of the various speech organs form the phonetic and prosodic elements of speech; and (5) the *shaping process* in which the transfer and radiation functions of the vocal tract modify the spectrum of the waveforms generated within the vocal mechanism.

It is only through understanding the dynamics of speech production that descriptive phonetics can be established on a firm basis. To accomplish this objective all aspects of the innervation and articulatory processes of speech production must be investigated. These include the production of pulmonic and other driving pressures for generating speech sounds, the various articulatory and phonatory actions of the larynx, and the complex behaviors of the supraglottal articulators in forming and modifying sounds. There are several different ways in which sounds are generated in the vocal tract and some of these have been investigated more extensively than others. In the production of sound by the vocal mechanism more precise models are needed for most of the various manners of articulation,

including plosive, ejective, implosive, click, flap, trill, sibilant, fricative, and sonorant. A serious examination of the mechanism involved in any one of these manners of articulation shows it to be a complex subject requiring both extensive empirical and theoretical research.

The fact that many different phone types may serve as allophones of the same phoneme has been repeatedly perplexing in the interpretation of the speech signal. Even more perplexing is the fact that phones of the same phone type may serve as allophones of different phonemes. The above facts make it quite unrealistic to search for either acoustical or physiological invariants of phonemes. It also emphasizes the importance of beginning with basic phonetic data and studying their organization in the hierarchy of phonological units in the interpretation of spoken language. Though speech is now generally considered to be highly complex, there is nothing particularly metaphysical about its structure and interpretation. The situation is simply that something more than elementary concepts of categorization are needed to interpret the phonological properties of speech in either the physiological or the acoustical form.

## Speech Acoustics

It has been increasingly recognized that at least some of the more peripheral processes of speech production must be well understood before the acoustical waves of speech can be interpreted in a fundamental way. Only a few years ago the primary problem in interpretating the speech wave was considered to be that of identifying its essential information bearing parameters. These parameters have now been reasonably well specified, although they certainly merit continued investigation. The truly difficult and important problem in acoustic phonetics is now before us. Specifically, we must learn how to interpret the dynamics of the acoustical parameters as discrete elements. In order to understand the acoustical parameters of speech, it was found necessary to investigate their origins in the physiology of speech production. In order to understand the dynamics of the acoustical speech signal, it is even more imperative that we investigate the dynamic aspects of the physiology of speech production.

## Speech Perception

Speech perception is a subdiscipline of speech science in which there has thus far been little basic progress. In fact, as yet the problems in this area have hardly been identified. Clearly, the observation of the responses of listeners to speech stimuli provides the primary experimental technique of the field.

Listening tests have been conducted in experimental studies of speech with a variety of objectives. The measurement of speech intelligibility is the traditional procedure with the most firmly established methods. Tests of speech intelligibility have been used under a variety of conditions and for a variety of purposes, but the chief use of these tests has been in the evaluation of communication systems.

Listening tests also serve a basic function in studies in experimental linguistics. In particular, listening tests provide a means of determining whether particular distinctions are actually present in the speech signal. One of the most effective procedures is to present a particular utterance and to ask listeners to respond by selecting from among a set of possible alternative orthographic equivalents.

Listening tests are also used frequently to conduct investigations in what is often termed " psycholinguistics ". These studies are generally directed at determining the relative complexities of phonological, lexical, or syntactic structures.

The most fundamental area of investigation in the field of speech perception, however, has received little attention thus far. It is well known in the study of other aspects of human perception that the background or context within which a stimulus appears may strongly affect the interpretation of the stimulus. There is certainly every reason to expect such effects in speech perception.

The most basic effects requiring investigation are at the phonetic and prosodic level. These effects should be distinguished from those resulting from lexical and syntactic influences. For example, it is highly probable that the perceived phonetic quality of a vowel phone type will be influenced on occasion by neighboring phone types. In some instances perceptual effects may extend over several elements in a string of phone types. These effects should be measurable in nonsense or foreign sequences, where lexical, syntactic, and semantic influences are minimal.

We do not know at present the extent of the problems in perceptual phonetics. Until such effects as those discussed above have been investigated more systematically, however, it is difficult to speculate on whether speech perception is intrinsically a significant field of research investigation. In order to investigate the field of speech perception in a serious way, it is quite obvious that the signal must be specified rigorously. Clearly, it is not possible to determine context effects of stimulus perception if the stimulus is ill-defined. The speech signal may be defined in speech wave types and values of the acoustical speech parameters as a function of time. Alternatively, it may be defined in articulatory and prosodic parameter values.

## Applied Speech Science

There are at least three major areas of important application for research in speech science. These areas are all based, or should be, in technically oriented experimental and theoretical investigations. When basic scientific approaches are employed in the study of applied problems in these areas, the knowledge gained in any one area is generally of value in understanding the others. Research in these areas is thus complementary, and the serious speech scientist usually has an interest in all of them.

*Phonological Description of Languages.* The first major area of the application of both experimental and theoretical knowledge in the field of speech science is the phonological description of specific spoken languages. In many respects this area

is basic to the two remaining applied areas to be discussed subsequently. The study of phonology is closely related to work in the field of linguistics. With the increased emphasis on grammar in linguistics, however, the experimental study of phonology is being left increasingly to the speech scientist. In fact, the reorientation in some areas of linguistics is much more suited to the study of the printed form of language than to the study of the spoken form.

There appears to be certain universals in spoken language which are independent of the type of language involved. Clearly, one of the basic responsibilities of those working in speech science is to identify the universals and to describe them in a formal way. It should be possible to specify experimental procedures for observing such universals. Attempts to develop a science of spoken language have been seriously handicapped by the acceptance of hypotheses and sometimes random speculations without experimental verification of any serious nature.

A scientist's description of the phonology of a specific language is clearly dependent upon his concept of phonological universals. If the description of the phonology of any particular language is to be valid, it must be based upon phenomena observed experimentally. According to this criterion there are few if any valid phonological descriptions of any languages at the present time. In fact, a complete phonological description of a language is a complicated structure with several major components.

*Speech Pathology.* The second major area of practical importance which is founded on the scientific study of speech is that of speech pathology. This area is closely related to medical science. A knowledge of the intricacies of the physiology of speech production is essential in effective work in speech pathology. Obviously, a reasonable knowledge of the phonological structure of the speech of the language in which the speech correction is conducted is essential to an effective program. This includes a knowledge of the phonetic, allophonic, phonemic, and prosodic aspects of the speech.

The speech pathologist deals with many physiological abnormalities and with numerous anomalies of speech production. He must have ways to identify and describe these, for only then can he know how to modify speech behavior toward a more normal pattern. A knowledge of the science of normal speech is certainly essential, but clearly it is not enough. An experimental and theoretical knowledge of the characteristics of the various disorders of speech is also required.

Some of the difficulties faced by the speech pathologist is indicated by present procedures in descriptive phonetics. The notations of descriptive phonetics have been developed primarily to characterize the great variety of formations which are distinctive in the various languages of the earth when spoken normally. While the basic parameters employed in a carefully formulated phonetic system are generally applicable to defective speech, they are thoroughly inadequate for describing many aspects of such speech. There are both new parameters and parameter values that must be added to accommodate the various kinds of formations found in defective speech. It is perhaps a commentary on the present state of

the field of speech science that speech pathologists are usually taught the phonetics of normal speech as if this symbolization were adequate for describing defective speech. The notation employed often constitutes a pseudophonemics rather than a phonetics. The concept of phone type and the numerous allophonic variations which often occur within phonemes are generally disregarded. Clearly, the field of descriptive phonetics should be extended, at least for speech correctionists, to encompass the kinds of pathological formations which they commonly encounter.

*Speech Communication.* The third major area of the application of information in the field of speech science is closely related to electronic engineering and especially to the information sciences. It extends from such established procedures as the measurement of speech intelligibility to the futuristic problems of language automation. Probably the current dominant problem in this area is that of the vocoder transmission of natural speech. The development of visual speech displays for spoken language instruction, for speech correction, and for teaching speech to the deaf are important problems within this field.

Automatic speech synthesis and automatic speech recognition are the primary constituents of the field of speech automation. The automatic translation from one spoken language to another is a problem in language automation so complex that it hardly merits notice at this time.

A valid phonological description of the language in which automatic speech synthesis and automatic speech recognition are to be achieved is an obvious necessity for successful speech automation. The combination of automatic speech synthesis and automatic speech recognition would provide a means of speech communication with a computer, which is a problem that will undoubtedly receive intensive consideration during the coming years.

*Summary.* I have considered the nature of speech science as an academic discipline. The primary basic and applied subdisciplines which are associated with the field have also been discussed. The basic areas encompass the various processes of the physiology of speech production, the acoustical characteristics of speech, including the speech wave types and the acoustical parameters of speech, and the processes of speech perception. Applied problems include the phonological description of languages, speech pathology, and speech communication. All of these subdisciplines, both basic and applied are extensive in nature and present innumerable and challenging research problems. These subdisciplines are all closely related and solid work in any one of them requires a substantial technical knowledge. In view of these considerations, it is indeed difficult to understand why the field has not yet reached the import and dimensions which make it recognizable as an independent discipline of academic instruction and research investigation. In the interim, there is an increasing number of research laboratories and institutes throughout the world devoted to the technical study of speech science and its applications. All of us must continue in some way to accommodate those students who come to this field with an enthusiasm for the challenges it presents and with a seriousness of purpose in the scientific investigation of its problems.